



SEQUENCE LISTING

<110> Dunn-Coleman, Nigel
Goedegebuur, Frits
Ward, Michael
Yao, Jian

<120> BGL5 Beta-Glucosidase and Nucleic Acids
Encoding the Same

<130> GC697

<140> US 10/026,140

<141> 2001-12-18

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1991

<212> DNA

<213> Trichoderma reesei

<400> 1

agccaggctcg	caaccagcag	cagcagcagc	agtacagaga	aatcaaccca	gatagctcaa	60
aatgcccagag	tcgctagctc	tgcccaacga	ctttgaatgg	ggcttcgcaa	cggccgccta	120
ccagatcgaa	ggcgccgtca	aagaagggtg	ccgcggcccc	tccatctggg	acacgtactg	180
ccacctggag	ccatcgcgca	ccaacggcgc	caacggcgat	gtggcttgcg	atcactacca	240
ccgctacgat	gaggactttg	atctcttgac	caagtacggc	gcaaaggcct	accgcttctc	300
cttgctgctg	tcgcggtatc	ttccccctcg	cggcaggctg	gatcccgtca	acgaggaggg	360
aattgagttt	tacagcaaac	tgattgacgc	cctggttgagg	cgggggtatc	cgccttggtg	420
gactttgtac	cactgggatc	tgccctcaggc	gcttcacgat	cgctatggag	gctgggtcaa	480
cgtggaagag	gtccagctgg	actttgagcg	gtatgcgagg	ttgtgctttg	aacgttttgg	540
ggaccgagtc	cagaactgga	tcacccatcaa	cgawccctgg	attcaggcca	tctatggata	600
tgccaccggc	agcaacgccc	cgggcaggag	cagcattaac	aagcactcca	ccgagggcaa	660
cactgccact	gagccgtggc	tcgctggaaa	ggcccagatc	atgagccatg	cccgcgcctg	720
ggcgcgtctac	agcagggact	ttcgccccctc	gcaaaagggc	cagatcggca	tctcgctcaa	780
cggcgactac	tatgagccct	gggacagcaa	tgagcctcgg	gacaaggagg	ctgctgagcg	840
acggatggaa	tttcacattg	gctggtttgc	caatcccctc	ttcttgaaga	aggactatcc	900
agagagcatg	aagaagcagc	tgggcgagag	gcttcacggc	ctcactcccg	cggactttgc	960
catectcaat	gccggagaga	ccgacttcta	cggcatgaat	tactacacat	cccagttcgc	1020
gcgccacctc	gacggtcccg	tccccgagac	ggactatctc	ggcgccatcc	atgagcacca	1080
ggagaataag	gacggcagcc	ccgttggcga	ggagagcggc	ctcgccctggc	tgcgctcctg	1140
cccggacatg	ttccggaagc	atctcgcccc	gggtgtacggc	ctgtacggca	agcccatcta	1200
catcaccgag	aacggatgcc	cgtgcccttg	agaggagaac	atgacgtgcg	aggaggccgt	1260
caacgacccc	ttccgcctcc	gstactttga	ctcgcaactg	gactcgattt	ccaaggccat	1320
taccacggac	ggcgtcgtcg	tcaaggggta	ctttgcgtgg	gcgttgctcg	ataacttgga	1380
atggtcagat	ggctacggac	ccagattcgg	cgtcacgttc	acagactaca	ccaccctcaa	1440
gcgcacgccc	aagaagtctg	ccctggctct	caaggacatg	tttgcgggcc	ggcagagggg	1500
taaagtggcg	gcataaagaa	agggaaattt	cttcttgcct	tcagcctcta	tgcatcttcc	1560
tctctctttt	tccctccctc	cccttgctcc	tctctctcta	cctctcatat	tccctctata	1620
ccccccgctt	cttctcatga	ccccatgctc	cttgcccttg	gccccctctc	gtcgaattct	1680
gcctcttctc	acgtcttatg	cgtctgttta	cttgccctttt	ttttttttgt	ctctttctgt	1740
ctgtctgtct	gcctgtctat	gtgtacctat	ctggcccttc	gctcattggc	aacagatact	1800
agcacaagtt	caagcaagca	agcacgcaag	caagcaagca	agccagccat	caacggcatc	1860
aaagcccat	gttttagctc	atgttcacat	tgctatgtta	tctacatcag	ccattcacta	1920
ccaggcgaag	aggccacaga	gagtctcatc	gtcttacctg	tatatacgct	tttttaaaaa	1980
aaaaaaaaaa	a					1991

<210> 2
 <211> 484
 <212> PRT
 <213> Trichoderma reesei

<220>
 <221> VARIANT
 <222> (1)...(484)
 <223> Xaa = Any Amino Acid

<400> 2
 Met Pro Glu Ser Leu Ala Leu Pro Asn Asp Phe Glu Trp Gly Phe Ala
 1 5 10 15
 Thr Ala Ala Tyr Gln Ile Glu Gly Ala Val Lys Glu Gly Gly Arg Gly
 20 25 30
 Pro Ser Ile Trp Asp Thr Tyr Cys His Leu Glu Pro Ser Arg Thr Asn
 35 40 45
 Gly Ala Asn Gly Asp Val Ala Cys Asp His Tyr His Arg Tyr Asp Glu
 50 55 60
 Asp Phe Asp Leu Leu Thr Lys Tyr Gly Ala Lys Ala Tyr Arg Phe Ser
 65 70 75 80
 Leu Ser Trp Ser Arg Ile Ile Pro Leu Gly Gly Arg Leu Asp Pro Val
 85 90 95
 Asn Glu Glu Gly Ile Glu Phe Tyr Ser Lys Leu Ile Asp Ala Leu Leu
 100 105 110
 Arg Arg Gly Ile Thr Pro Trp Val Thr Leu Tyr His Trp Asp Leu Pro
 115 120 125
 Gln Ala Leu His Asp Arg Tyr Gly Gly Trp Leu Asn Val Glu Glu Val
 130 135 140
 Gln Leu Asp Phe Glu Arg Tyr Ala Arg Leu Cys Phe Glu Arg Phe Gly
 145 150 155 160
 Asp Arg Val Gln Asn Trp Ile Thr Ile Asn Xaa Pro Trp Ile Gln Ala
 165 170 175
 Ile Tyr Gly Tyr Ala Thr Gly Ser Asn Ala Pro Gly Arg Ser Ser Ile
 180 185 190
 Asn Lys His Ser Thr Glu Gly Asn Thr Ala Thr Glu Pro Trp Leu Ala
 195 200 205
 Gly Lys Ala Gln Ile Met Ser His Ala Arg Ala Val Ala Val Tyr Ser
 210 215 220
 Arg Asp Phe Arg Pro Ser Gln Lys Gly Gln Ile Gly Ile Ser Leu Asn
 225 230 235 240
 Gly Asp Tyr Tyr Glu Pro Trp Asp Ser Asn Glu Pro Arg Asp Lys Glu
 245 250 255
 Ala Ala Glu Arg Arg Met Glu Phe His Ile Gly Trp Phe Ala Asn Pro
 260 265 270
 Ile Phe Leu Lys Lys Asp Tyr Pro Glu Ser Met Lys Lys Gln Leu Gly
 275 280 285
 Glu Arg Leu Pro Ala Leu Thr Pro Ala Asp Phe Ala Ile Leu Asn Ala
 290 295 300
 Gly Glu Thr Asp Phe Tyr Gly Met Asn Tyr Tyr Thr Ser Gln Phe Ala
 305 310 315 320
 Arg His Leu Asp Gly Pro Val Pro Glu Thr Asp Tyr Leu Gly Ala Ile
 325 330 335
 His Glu His Gln Glu Asn Lys Asp Gly Ser Pro Val Gly Glu Glu Ser
 340 345 350
 Gly Leu Ala Trp Leu Arg Ser Cys Pro Asp Met Phe Arg Lys His Leu
 355 360 365
 Ala Arg Val Tyr Gly Leu Tyr Gly Lys Pro Ile Tyr Ile Thr Glu Asn
 370 375 380

Gly	Cys	Pro	Cys	Pro	Gly	Glu	Glu	Asn	Met	Thr	Cys	Glu	Glu	Ala	Val
385					390					395					400
Asn	Asp	Pro	Phe	Arg	Ile	Arg	Tyr	Phe	Asp	Ser	His	Leu	Asp	Ser	Ile
				405					410					415	
Ser	Lys	Ala	Ile	Thr	Gln	Asp	Gly	Val	Val	Val	Lys	Gly	Tyr	Phe	Ala
			420					425					430		
Trp	Ala	Leu	Leu	Asp	Asn	Leu	Glu	Trp	Ser	Asp	Gly	Tyr	Gly	Pro	Arg
		435					440					445			
Phe	Gly	Val	Thr	Phe	Thr	Asp	Tyr	Thr	Thr	Leu	Lys	Arg	Thr	Pro	Lys
	450					455					460				
Lys	Ser	Ala	Leu	Val	Leu	Lys	Asp	Met	Phe	Ala	Ala	Arg	Gln	Arg	Val
465					470					475					480
Lys	Val	Ala	Ala												

<210> 3

<211> 1455

<212> DNA

<213> Trichoderma reesei

<400> 3

atgcccgagt	cgctagctct	gcccacgac	tttgaatggg	gcttcgcaac	ggccgcctac	60
cagatcgaag	gcgccgtcaa	agaaggtggc	cgcggcccg	ccatctggga	cacgtactgc	120
cacctggagc	catcgcgcac	caacggcgcc	aacggcgatg	tggcttgcca	tcactaccac	180
cgctacgatg	aggactttga	tctcttgacc	aagtacggcg	caaaggccta	ccgcttctcc	240
ttgtcgtggt	cgcggtatcat	tcccctcggc	ggcaggctgg	atcccgtcaa	cgaggaggga	300
attgagtttt	acagcaaact	gattgacgcc	ctgttgaggc	gggggtatcac	gccttgggtg	360
actttgtacc	actgggatct	gcctcaggcg	cttcacgatc	gctatggagg	ctggctcaac	420
gtggaagagg	tccagctgga	ctttgagcgg	tatgcgaggt	tgtgctttga	acgttttggg	480
gaccgagtcc	agaactggat	caccatcaac	gawccctgga	ttcaggccat	ctatggatat	540
gccaccggca	gcaacgcccc	gggcaggagc	agcattaaca	agcactccac	cgaggggcaac	600
actgccactg	agccgtggct	cgctggaaa	gcccagatca	tgagccatgc	ccgcgccgtg	660
gccgtctaca	gcagggaact	tcgccccctc	caaaagggcc	agatcggcac	ctcgctcaac	720
ggcgactact	atgagccctg	ggacagcaat	gagcctcggg	acaaggaggc	tgctgagcga	780
cggatggaat	ttcacattgg	ctggtttgcc	aatcccatct	tcttgaagaa	ggactatcca	840
gagagcatga	agaagcagct	gggcgagagg	cttcacagccc	tcactcccgc	ggactttgcc	900
atcctcaatg	ccggagagac	cgacttctac	ggcatgaatt	actacacatc	ccagttcgcg	960
cgccacctag	acgggtcccg	ccccgagacg	gactatctcg	gcgccatcca	tgagcaccag	1020
gagaataagg	acggcagccc	cgttggcgag	gagagcggcc	tcgcctggct	gcgctcctgc	1080
ccggacatgt	tccggaagca	tctcgcccgg	gtgtacggcc	tgtacggcaa	gcccatctac	1140
atcaccgaga	acggatgccc	gtgccctgga	gaggagaaca	tgacgtgcga	ggaggccgtc	1200
aacgaccctt	tccgcatccg	stactttgac	tcgcacttgg	actcgatttc	caaggccatt	1260
acccaggacg	gcgtcgtcgt	caaggggtac	tttgcggtgg	cgttgctcga	taacttggaa	1320
tggtcagatg	gctacggacc	cagattcggc	gtcacgttca	cagactacac	caccctcaag	1380
cgcacgcca	agaagtctgc	cctgggtctc	aaggacatgt	ttgcggcccc	gcagagggtt	1440
aaagtggcgg	cataa					1455